



Why 80% Accuracy Doesn't Always Mean a Child Has Mastered a Skill

By Dr. Peter Beddow, BCBA-D



Understanding percentages, sample sizes, and what data really tell us about learning.

In applied behavior analysis, we frequently use percentages to describe how accurately a child performs a skill. We might say that a child identified pictures with 80 percent accuracy, followed directions with 90 percent accuracy, or completed a task independently on 100 percent of opportunities. Those numbers are useful, but they can also be misleading if we do not consider how they were obtained.

I was recently reviewing a data sheet with a therapist who had recorded five trials of a motor imitation program. The child responded correctly on four trials and incorrectly on one. That gave us a score of 80 percent.

Mathematically, the score was correct. But I asked the therapist whether she thought those five trials really told us enough about what the child knew. Four correct responses out of five do not necessarily mean the child has mastered the skill.

A percentage describes the responses that were recorded. It does not automatically tell us whether the sample was large enough, representative enough, or independent enough to support a mastery decision.

Five Trials Can Produce a Very Different Picture From Ten

Let's imagine that we are teaching a child to follow two motor instructions: jump and spin. We present five trials, and the child responds correctly on four of them. That is 80 percent accuracy.

Now suppose we present five additional trials. This time, the child responds correctly on only two. Across all ten trials, the child has responded correctly six times, giving us an overall score of 60 percent.

The original 80 percent was not wrong. It accurately described what happened during those first five trials. But it did not provide a complete picture of the child's performance across a larger sample.

This is why the number of opportunities matters. A percentage based on five trials may be more sensitive to one or two responses than a percentage based on a larger sample. We need to consider whether the sample adequately represents the skill we are trying to measure.

The number of trials should be appropriate for the skill, the child, and the measurement plan. Five opportunities may be useful for one purpose and insufficient for another. What matters is that the team understands what the data can and cannot support.

We should also pay attention to when the trials occurred. A child may perform differently at the beginning and end of a session, in a quiet room and a busy classroom, or with familiar and unfamiliar materials. A percentage without information about the conditions may leave out important parts of the story.

The Same Percentage Can Hide Different Problems

There is another issue with percentages. Sometimes an overall score combines performance on several different targets, and that can make it difficult to identify what the child actually knows.

Suppose we are teaching jump and spin together. The child completes ten trials and responds correctly on eight. That is 80 percent accuracy, but what if all the incorrect responses occurred when we asked the child to spin? The child may have mastered jumping while still needing additional instruction on spinning.

An overall percentage does not necessarily reveal that pattern. We may need to examine performance by target or record which responses were incorrect. That additional information can help us make better instructional decisions.

The goal is not simply to know how many responses were correct. We want to understand which skills are strong and which ones still need teaching.

This applies to other kinds of targets as well. A child may identify familiar objects accurately but struggle with less familiar examples. They may follow one-step directions but have difficulty with two-step directions. They may answer questions about a favorite topic but not about a new story. Combining all of those responses into one number could hide the distinction that matters most for instruction.

One Correct Response Does Not Always Demonstrate Learning

Suppose I place two pictures on the table and ask a child to touch the car. The child touches the correct picture. Can I conclude that the child knows what a car is?

Not necessarily. With two equally likely choices, a child who is guessing has a 50 percent chance of selecting the correct picture. One correct response does not provide strong evidence that the child has learned the discrimination.

That is why we often present multiple opportunities and rearrange the materials. We want to see whether the child consistently selects the correct picture across different positions and combinations. Repeated correct responding under appropriate conditions gives us stronger evidence of learning.

The same principle applies to many other skills. We need to consider how likely a correct response would be without genuine understanding. If a child is choosing between several options, we should consider the number of choices, the arrangement of the materials, and whether one option is easier to select for reasons unrelated to the target.

A correct response is useful information. It becomes stronger evidence when it is repeated, independent, and observed under varied but relevant conditions.

Some Responses Provide Stronger Evidence Than Others

Now imagine that I hold up a phone and ask a child, “What’s this?” The child immediately says, “Phone.” Then I hold up a water bottle, and the child says, “Water bottle.”

Those responses may provide stronger evidence of object labeling than a single correct selection from a field of two. The child has produced specific words that correspond to the objects rather than choosing between two alternatives. If those are familiar, well-established skills, we may not need extensive additional teaching.

But we still need to be careful about declaring mastery of a new skill based on only a couple of responses. A child might have memorized particular examples or learned to respond under very specific conditions. Formal mastery decisions should be based on established criteria and representative performance over time.

The response format also matters. A child may correctly label a familiar object but not identify it when asked to select it from several options. These responses can be related without being identical. We should define the target clearly and interpret the data according to the skill we actually intended to teach.

Prompting Can Change What the Percentage Means

Another important question is whether the child responded independently. Suppose a therapist presents ten trials, and the child completes all ten responses correctly. That sounds like 100 percent accuracy.

But what if the therapist provided prompts on six of those trials? If we are measuring independent responding, the child was independent on only four out of ten opportunities. That is 40 percent independence, not 100 percent.

The prompted responses may have been useful teaching opportunities, but they do not demonstrate independent mastery. We need to make sure our data collection procedures distinguish between prompted and independent performance when that distinction is relevant to the learning target.

This does not mean prompted responses should be ignored. They can show that the child is learning with support and can help the team identify which prompts are useful. The important thing is to record them accurately so a prompted response is not mistaken for an independent one.

A data sheet may need separate codes for independent, prompted, incorrect, and no-response trials. The specific system can vary, but the meaning should be clear to everyone who uses it.

Mastery Criteria Need to Mean Something

Many instructional programs use mastery criteria such as 80 or 90 percent accuracy across a specified number of sessions. Those criteria can be useful because they provide a consistent standard for deciding when to move forward. But the percentage alone is not enough.

We also need to consider the number of trials, the consistency of performance, the presence of prompts, and whether the skill occurs across relevant conditions. A child who scores 80 percent on five trials during one session may not demonstrate the same level of performance the next day. A child who scores 80 percent across several sessions with different materials may provide much stronger evidence of learning.

Mastery criteria should be selected thoughtfully rather than treated as arbitrary numbers. The appropriate standard depends on the skill, the consequences of errors, and how the skill will be used.

For an important safety response, we may need stronger and more consistent evidence than for a skill where an occasional error has few consequences. A skill that must be used across many settings may need to be checked with different people and materials. A communication skill may need to be demonstrated when the child genuinely wants something, not only when an adult asks for a practice response.

The criterion should help the team make a responsible decision. Reaching a number is not the same as proving that a child can use a skill when it matters. A 2026 TABA conference presentation by Sarah Richling similarly questioned one-size-fits-all mastery criteria and proposed the more neutral term “performance criteria,” while noting that direct empirical research connecting mastery criteria to response maintenance remains limited [3].

Do Not Collect More Data Just to Get a Better Score

There is an important caution here. If a child scores poorly on the first five trials, we should not simply keep presenting trials until the percentage looks better. That would make our measurement inconsistent and could distort our interpretation of progress.

Instead, we should establish reasonable data collection procedures and follow them consistently. If additional trials are needed because the original sample was insufficient, we should record and interpret those trials transparently. We can also distinguish between formal measurement opportunities and additional teaching trials.

The goal is to obtain an accurate picture of performance, not to produce the highest possible percentage.

Sometimes a low score is useful information. It may show that the target is too difficult, that the child needs a different prompt, or that the teaching conditions are not supporting success. Changing the measurement simply to improve the number can hide the information we need in order to teach effectively.

Look Beyond the Number

When I review data, I want to know more than whether the child reached 80 percent. I want to know what happened during the trials, whether the responses were independent, and whether the child demonstrated the skill consistently. I also want to know whether the performance matches what therapists are observing during everyday activities.

If the data and observations do not seem to agree, that is a reason to investigate. Maybe we need more information, or maybe there is something about the teaching arrangement that we have not recognized. The percentage is a starting point for understanding performance, not the end of the conversation.

We might examine the individual targets, review the prompts, vary the materials, observe another session, or check whether the skill appears in a natural setting. These steps can help us understand why a number looks the way it does.

Data are supposed to help us make better decisions about teaching. Sometimes the most important question is not, “What percentage did the child get?”

It is, “What does that percentage actually tell us?”

References

- [1] Kazdin, A. E. (2011). *Single-Case Research Designs: Methods for Clinical and Applied Settings* (2nd ed.). Oxford University Press. [<https://global.oup.com/academic/product/single-case-research-designs-9780190079971>] [2] Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied Behavior Analysis* (3rd ed.). Pearson. [<https://www.pearson.com/en-us/subject-catalog/p/applied-behavior-analysis/P200000001665>] [3] Richling, S. (2026, September 17). The “Mastery Criteria” Myth: One-Size Does NOT Fit All in ABA Service Delivery [Conference presentation]. 29th Annual Conference of the Tennessee Association for Behavior Analysis (TABA), Nashville, TN.
-

Continue learning

Open the companion resource and choose the next practical step.

<https://accessiblebehavior.com/go/why-80-accuracy-doesn-t-always-mean-a-child-has-mastered-a-skill>

About the author

Teacher at heart, behavior analyst, and writer, Dr. Peter Beddow has spent more than two decades helping children, families, and school teams navigate autism, behavior, trauma, and learning differences with clarity, structure, and hope. He studied English Literature and Creative Writing at Middlebury College and earned a PhD in Special Education from Vanderbilt University's Peabody College, where his work focused on fairness, learning, and access in assessment. Peter has worked as a teacher, special educator, clinical director, consultant, and board certified behavior analyst in homes, schools, and clinics. He founded Accessible Hope, a Nashville-area practice supporting diverse families, and writes and teaches practical, humane approaches to behavior and learning.

Scope and safety

This article provides general educational information. It is not individualized medical, behavioral, mental-health, legal, or emergency advice. Individual support should consider the learner's communication, health, safety, disability-related needs, culture, and context.

How to cite this resource

Beddow, P. (2026). Why 80% Accuracy Doesn't Always Mean a Child Has Mastered a Skill. Accessible Behavior.
<https://accessiblebehavior.com/go/why-80-accuracy-doesn-t-always-mean-a-child-has-mastered-a-skill>

Copyright © 2026 Accessible Behavior. All rights reserved.

Published article: <https://accessiblebehavior.com/resources/why-80-accuracy-doesn-t-always-mean-a-child-has-mastered-a-skill/>